Shashwat Raj

Bengaluru / Delhi, India

shashwatraj98765@gmail.com | +91 89863 28340 | linkedin/iamsrajj | github/iamsrajj | iamsrajj.infean.in

Summary

M.Tech. student in Machine Intelligence at IIT Delhi, with a B.Tech. in Information Technology (AI & ML) from NITK Surathkal. Experienced in building production-grade AI systems, scalable backend APIs, and LLM-powered applications for real-world use cases. Skilled in Python, Generative AI, Retrieval-Augmented Generation (RAG), prompt engineering, and context-aware system design. Passionate about creating intelligent, reliable, and impactful AI products.

Education

Indian Institute of Technology (IIT) Delhi	**Expected 2027**
M.Tech. in Machine Intelligence	
National Institute of Technology Karnataka (NITK), Surathkal	**2026**
B.Tech. in Information Technology (AI & ML)	CGPA: **8.9/10**
Bal Bhawan Public School, Bhopal	**2022**
Higher Secondary (Class XII)	Percentage: **87%**

Work Experience

Co-Founder / GenAI Engineer, AgriDoot by NovosEdge Pvt. Ltd.　　　　**Jan 2026 – Present**
India

- Co-built an AI-powered agricultural intelligence platform delivering real-time advisory and personalized recommendations to farmers.
- Engineered a context-aware GenAI chatbot (*Gyan*) using LLMs for domain-specific conversational assistance.
- Integrated weather data, mandi rates, and AI-driven insights into a unified decision-support system.
- Architected modular Flask-based APIs for scalable backend deployment and personalized user workflows.
- Led AI system design, backend development, deployment workflows, and product scaling strategies.

Project Intern, Microsoft　　　　**Oct 2025 – Dec 2025**
Remote

- Contributed to AI/ML-based solutions focused on real-world product and engineering use cases.
- Supported model development, experimentation, and scalable system design initiatives.
- Collaborated with cross-functional teams to build efficient, reliable, and production-oriented solutions.

Researcher, IIT Delhi　　　　**Feb 2025 – Jul 2025**
On-site

- Conducted research in Artificial Intelligence and Machine Learning domains with a focus on practical applications.
- Worked on model development, experimentation pipelines, and performance evaluation.
- Explored advanced AI techniques and their applicability to real-world problem settings.

Software Developer Intern, Wipro Ltd.　　　　**Sep 2023 – Dec 2023**

- Developed data-driven solutions using Python and analytical workflows for business use cases.
- Performed data analysis to extract insights and improve process efficiency.
- Enhanced workflows and user experience through system optimization initiatives.

Software Developer Intern, Prudent Systems Pvt. Ltd.　　　　**Mar 2023 – Aug 2023**

- Built custom workflows, validation rules, and backend logic for business applications.
- Developed automated email systems and notification services using Python.
- Supported internal data analysis and model training tasks across projects.

Project Developer, Hackster.io (Part-time)　　　　**Jan 2021 – Dec 2022**
Remote

- Developed IoT-based and AI-driven prototypes as part of research and development initiatives.
- Built solutions using Arduino, Python, and web technologies.
- Integrated hardware systems with intelligent software workflows for prototype development.

Skills

Languages: Python, HTML, CSS
AI/ML: PyTorch, TensorFlow, Scikit-learn, CNN, RNN

Generative AI: Large Language Models (LLMs), Prompt Engineering, Retrieval-Augmented Generation (RAG), Context Management, Memory Systems
Backend: Flask, REST API Development, System Design, Scalable APIs
Databases: MySQL, PostgreSQL, FAISS, Pinecone
Tools & Infrastructure: Google Cloud Platform (GCP), Git, Linux (Ubuntu), Apache, Nginx
AI Engineering Practices: Model Evaluation, Prompt Testing, Latency Optimization, Cost Optimization, Error Handling, Hallucination Mitigation
Additional Skills: Problem Solving, Debugging, Communication, Collaboration, Team Leadership, Product Thinking

Projects

AgriDoot – AI-Powered Agricultural Intelligence Platform

- Co-developed a production-grade AI platform providing real-time advisory and decision support to farmers.

- Integrated weather data, mandi rates, and AI-driven recommendations into a unified intelligence system.

- Built modular Flask APIs with scalable architecture and API key-based authentication.

- Enabled chatbot assistance, disease detection, and personalized agricultural insights.

Gyan Advanced – GenAI System with Context Injection

- Designed an advanced LLM pipeline using Drona and Gemma with dynamic system prompts.

- Integrated contextual user farm data (GIS, crop, location) for personalized response generation.

- Implemented conversation history tracking for follow-up query handling and continuity.

- Improved response relevance using structured prompt engineering and context injection strategies.

Advanced RAG System with Hybrid Search

- Built a Retrieval-Augmented Generation (RAG) pipeline combining vector search and BM25-based keyword retrieval.

- Improved answer accuracy and reduced hallucinations through hybrid retrieval and re-ranking.

- Designed document chunking, embedding generation, and optimized query pipelines.

- Integrated FAISS / Pinecone with context filtering and prompt optimization techniques.

Project K – Personalized LLM Companion for Conversational AI

- Built a fine-tuned LLaMA-based conversational AI system for personalized and human-like interactions.

- Enabled long-term memory and context retention to support continuous evolving conversations.

- Designed prompt engineering workflows to improve conversational relevance and continuity.

- Implemented safety layers to reduce harmful, unsafe, or biased outputs.

AgriDoot Developer Console – API Platform

- Built a developer platform for accessing AI APIs including Gyan Chat, Vision, Weather, and Mandi services.

- Designed modular API architecture with categorized endpoints and scalable backend workflows.

- Implemented authentication, logging, and deployment pipelines using systemd and Nginx.

- Created structured API documentation to improve developer usability and integration.

Foliage Detect – AI-Based Plant Disease Detection System

- Developed an AI-powered image classification system for plant disease detection.

- Trained deep learning models using CNN architectures for leaf image analysis.

- Built an API pipeline for real-time prediction and structured result storage.

- Improved usability through a database-backed disease tracking workflow.

Data Analyzer – Automated Data Analysis & Visualization Tool

- Built a Python-based system to process, clean, and analyze CSV datasets.

- Implemented customizable visualizations and automated statistical insight generation.

- Enabled flexible data transformation workflows for varied datasets and use cases.

- Designed the system for reusable analysis across multiple domains.

Publications

Optimizing IoT Communication with LoRa
Research Publication – Available on ResearchGate
ResearchGate Link

AI Demystified – A Beginner's Handbook
Author – Published on Amazon
Amazon Link

Awards

- **Best Project of the Year (2025–2026)** – Recognized for outstanding innovation and implementation.
- **Gold Medalist (2025–2026)** – Awarded for academic excellence.

Certifications

- Complete Generative AI Course with LangChain and Hugging Face – Udemy
- Artificial Intelligence A-Z 2025: Agentic AI, Gen AI, and RL – Udemy
- Machine Learning A-Z: AI, Python & R + ChatGPT Prize [2025] – Udemy
- The Complete Python Bootcamp From Zero to Hero in Python – Udemy
- Advanced Certification in Data Science and AI – IIT Madras / Intellipaat
- Published Author Certificate – Notion Press / Amazon Publishing
- Certificate of Achievement – MANIT Bhopal
- Certificate of Achievement – NIT Karnataka
- Solve for Tomorrow 2023 – Samsung
- Remote Sensing & GIS for Environmental Studies – ISRO / IIRS Dehradun